

Programming an Interpreter Using Molecular Dynamics^{*}

J.A. Bergstra^{1,2} and C.A. Middelburg¹

¹ Programming Research Group, University of Amsterdam,
P.O. Box 41882, 1009 DB Amsterdam, the Netherlands

² Department of Philosophy, Utrecht University,
P.O. Box 80126, 3508 TC Utrecht, the Netherlands
J.A.Bergstra@uva.nl, C.A.Middelburg@uva.nl

Abstract. PGA (ProGram Algebra) is an algebra of programs which concerns programs in their simplest form: sequences of instructions. Molecular dynamics is a simple model of computation developed in the setting of PGA, which bears on the use of dynamic data structures in programming. We consider the programming of an interpreter for a program notation that is close to existing assembly languages using PGA with the primitives of molecular dynamics as basic instructions. It happens that, although primarily meant for explaining programming language features relating to the use of dynamic data structures, the collection of primitives of molecular dynamics in itself is suited to our programming wants.

Keywords: interpreter, program algebra, molecular dynamics, thread algebra, action transforming service.

1998 ACM Computing Classification: D.1.4, D.3.1, D.3.4, F.1.1, F.3.2.

1 Introduction

In this paper, we consider the programming of an interpreter for a program notation that is close to existing assembly languages using PGA (ProGram Algebra). With that we carry on the line of research with which a start was made in [3]. The object pursued with that line of research is the development of a theoretical understanding of possible forms of sequential programs, starting from the simplest form of sequential programs, and associated ways of programming. The view is taken that sequential programs in the simplest form are sequences of instructions. PGA, an algebra of programs in which programs are looked upon as sequences of instructions, is taken for the basis of the development aimed at. Until now, the research was rather focussed on the development of a theoretical understanding of possible forms of sequential programs. The work presented in the current paper is primarily concerned with programming using the simplest form of sequential programs and molecular dynamics.

^{*} This research was partly carried out in the framework of the Jacquard-project Symbiosis, which is funded by the Netherlands Organisation for Scientific Research (NWO).

For the programming of the interpreter, we use PGA with the primitives of molecular dynamics as basic instructions. Molecular dynamics is a simple model of computation bearing on the use of dynamic data structures in programming. In this model, states of computations resemble collections of molecules composed of atoms and computations take place by means of actions which transform the structure of molecules like in chemical reactions. Molecular dynamics has been developed mainly in the setting of PGA. The model introduced in the current paper is a small expansion of the model that was first described informally in [1] and formally in [5].

In the line of research carried on, the view is taken that the behaviours exhibited by sequential programs on execution are threads as considered in BTA (Basic Thread Algebra).³ A thread proceeds by doing steps in a sequential fashion. A thread may do certain steps for having itself affected by some service or affecting some service. In the current paper, we will use a generalization of the use mechanism introduced in [8] and a complementary mechanism of that mechanism for such interactions between threads that are behaviours exhibited by sequential programs on execution and services that deal with steps that relate to molecular dynamics. A slightly different generalization of the use mechanism from [8] is introduced in [6] under the name action transforming thread-service composition.

A hierarchy of program notations rooted in PGA is introduced in [3]. In the current paper, we consider the programming of an interpreter for one program notation that belongs to this hierarchy. The program notation in question, called PGLD (ProGramming Language D), is a very simple program notation which is close to existing assembly languages. The hierarchy also includes a program notation, called PGLS (ProGramming Language for Structured programming), that supports structured programming by offering a rendering of conditional and loop constructs instead of (unstructured) jump instructions. Each PGLS program can be translated into a semantically equivalent PGLD program. In [7], a variant of PGLD with indirect jump instructions is introduced. In the current paper, we show how to adapt the interpreter for PGLD to the presence of indirect jump instructions.

This paper is organized as follows. First, we review BTA, PGA, and PGLD (Sections 2, 3, and 4). Next, we extend BTA with the mechanisms for interaction between threads and services used, introduce a state-based approach to describe services, and give a state-based description of services for molecular dynamics (Sections 5, 6, and 7). Following this, we give PGA programs for creating representations of PGLD programs by molecules and a PGA program for interpreting those representations (Sections 8 and 9). After that, we introduce the variant of PGLD with indirect jump instructions and adapt the above-mentioned PGA programs to the presence of indirect jump instructions (Sections 10 and 11). Finally, we make some concluding remarks (Section 12).

³ In [3], basic thread algebra is introduced under the name basic polarized process algebra. Prompted by the development of thread algebra [8], which is a design on top of it, basic polarized process algebra has been renamed to basic thread algebra.

Table 1. Axiom of BTA

$$\frac{x \trianglelefteq \mathbf{tau} \trianglerighteq y = x \trianglelefteq \mathbf{tau} \trianglerighteq x}{\text{T1}}$$

2 Basic Thread Algebra

In this section, we review BTA, a form of process algebra which is tailored to the description of the behaviour of deterministic sequential programs under execution. The behaviours concerned are called *threads*.

In BTA, it is assumed that there is a fixed but arbitrary finite set of *basic actions* \mathcal{A} with $\mathbf{tau} \notin \mathcal{A}$. We write $\mathcal{A}_{\mathbf{tau}}$ for $\mathcal{A} \cup \{\mathbf{tau}\}$. The members of $\mathcal{A}_{\mathbf{tau}}$ are referred to as *actions*.

The intuition is that each basic action performed by a thread is taken as a command to be processed by a service provided by the execution environment of the thread. The processing of a command may involve a change of state of the service concerned. At completion of the processing of the command, the service produces a reply value. This reply is either **T** or **F** and is returned to the thread concerned.

Although BTA is one-sorted, we make this sort explicit. The reason for this is that we will extend BTA with an additional sort in Section 5.

The algebraic theory BTA has one sort: the sort **T** of *threads*. To build terms of sort **T**, BTA has the following constants and operators:

- the *deadlock* constant $\mathbf{D} : \mathbf{T}$;
- the *termination* constant $\mathbf{S} : \mathbf{T}$;
- for each $a \in \mathcal{A}_{\mathbf{tau}}$, the binary *postconditional composition* operator $- \trianglelefteq a \trianglerighteq - : \mathbf{T} \times \mathbf{T} \rightarrow \mathbf{T}$.

Terms of sort **T** are built as usual (see e.g. [13, 14]). Throughout the paper, we assume that there are infinitely many variables of sort **T**, including x, y, z .

We use infix notation for postconditional composition. We introduce *action prefixing* as an abbreviation: $a \circ p$, where p is a term of sort **T**, abbreviates $p \trianglelefteq a \trianglerighteq p$.

Let p and q be closed terms of sort **T** and $a \in \mathcal{A}_{\mathbf{tau}}$. Then $p \trianglelefteq a \trianglerighteq q$ will perform action a , and after that proceed as p if the processing of a leads to the reply **T** (called a positive reply), and proceed as q if the processing of a leads to the reply **F** (called a negative reply). The action \mathbf{tau} plays a special role. It is a concrete internal action: performing \mathbf{tau} will never lead to a state change and always lead to a positive reply, but notwithstanding all that its presence matters.

BTA has only one axiom. This axiom is given in Table 1. Using the abbreviation introduced above, axiom T1 can be written as follows: $x \trianglelefteq \mathbf{tau} \trianglerighteq y = \mathbf{tau} \circ x$.

Each closed BTA term of sort **T** denotes a finite thread, i.e. a thread of which the length of the sequences of actions that it can perform is bounded. Guarded recursive specifications give rise to infinite threads.

Table 2. Axioms for guarded recursion

$\langle X E \rangle = \langle t_X E \rangle$	if $X = t_X \in E$	RDP
$E \Rightarrow X = \langle X E \rangle$	if $X \in V(E)$	RSP

A *guarded recursive specification* over BTA is a set of recursion equations $E = \{X = t_X \mid X \in V\}$, where V is a set of variables of sort \mathbf{T} and each t_X is a term of the form D, S or $t \triangleleft a \triangleright t'$ with t and t' BTA terms of sort \mathbf{T} that contain only variables from V . We write $V(E)$ for the set of all variables that occur on the left-hand side of an equation in E . We are only interested in models of BTA in which guarded recursive specifications have unique solutions, such as the projective limit model of BTA presented in [2]. A thread that is the solution of a finite guarded recursive specification over BTA is called a *finite-state* thread.

We extend BTA with guarded recursion by adding constants for solutions of guarded recursive specifications and axioms concerning these additional constants. For each guarded recursive specification E and each $X \in V(E)$, we add a constant of sort \mathbf{T} standing for the unique solution of E for X to the constants of BTA. The constant standing for the unique solution of E for X is denoted by $\langle X|E \rangle$. Moreover, we add the axioms for guarded recursion given in Table 2 to BTA, where we write $\langle t_X|E \rangle$ for t_X with, for all $Y \in V(E)$, all occurrences of Y in t_X replaced by $\langle Y|E \rangle$. In this table, X , t_X and E stand for an arbitrary variable of sort \mathbf{T} , an arbitrary BTA term of sort \mathbf{T} and an arbitrary guarded recursive specification over BTA, respectively. Side conditions are added to restrict the variables, terms and guarded recursive specifications for which X , t_X and E stand. The equations $\langle X|E \rangle = \langle t_X|E \rangle$ for a fixed E express that the constants $\langle X|E \rangle$ make up a solution of E . The conditional equations $E \Rightarrow X = \langle X|E \rangle$ express that this solution is the only one.

We will use the following abbreviation: a^ω , where $a \in \mathcal{A}_{\text{tau}}$, abbreviates $\langle X|\{X = a \circ X\} \rangle$.

We will write BTA+REC for BTA extended with the constants for solutions of guarded recursive specifications and axioms RDP and RSP.

In [4], we show that the threads considered in BTA+REC can be viewed as processes that are definable over ACP [10].

Closed terms of sort \mathbf{T} from the language of BTA+REC that denote the same infinite thread cannot always be proved equal by means of the axioms of BTA+REC. We introduce the approximation induction principle to remedy this. The approximation induction principle, AIP in short, is based on the view that two threads are identical if their approximations up to any finite depth are identical. The approximation up to depth n of a thread is obtained by cutting it off after performing a sequence of actions of length n .

AIP is the infinitary conditional equation given in Table 3. Here, following [3], approximation of depth n is phrased in terms of a unary *projection* operator $\pi_n : \mathbf{T} \rightarrow \mathbf{T}$. The axioms for the projection operators are given in Table 4. In this table, a stands for an arbitrary member of \mathcal{A}_{tau} .

Table 3. Approximation induction principle

$$\overline{\bigwedge_{n \geq 0} \pi_n(x) = \pi_n(y) \Rightarrow x = y} \text{ AIP}$$

Table 4. Axioms for projection operators

$\pi_0(x) = \mathsf{D}$	P0
$\pi_{n+1}(\mathsf{S}) = \mathsf{S}$	P1
$\pi_{n+1}(\mathsf{D}) = \mathsf{D}$	P2
$\pi_{n+1}(x \sqsubseteq a \sqsupseteq y) = \pi_n(x) \sqsubseteq a \sqsupseteq \pi_n(y)$	P3

We will write BTA+REC+AIP for BTA+REC extended with the projection operators and the axioms from Tables 3 and 4.

3 Program Algebra

In this section, we review PGA, an algebra of sequential programs based on the idea that sequential programs are in essence sequences of instructions. PGA provides a program notation for finite-state threads.

In PGA, it is assumed that there is a fixed but arbitrary finite set \mathfrak{A} of *basic instructions*. PGA has the following *primitive instructions*:

- for each $a \in \mathfrak{A}$, a *plain basic instruction* a ;
- for each $a \in \mathfrak{A}$, a *positive test instruction* $+a$;
- for each $a \in \mathfrak{A}$, a *negative test instruction* $-a$;
- for each $l \in \mathbb{N}$, a *forward jump instruction* $\#l$;
- a *termination instruction* $!$.

We write \mathfrak{I} for the set of all primitive instructions.

The intuition is that the execution of a basic instruction a may modify a state and produces T or F at its completion. In the case of a positive test instruction $+a$, basic instruction a is executed and execution proceeds with the next primitive instruction if T is produced and otherwise the next primitive instruction is skipped and execution proceeds with the primitive instruction following the skipped one. In the case where T is produced and there is not at least one subsequent primitive instruction and in the case where F is produced and there are not at least two subsequent primitive instructions, deadlock occurs. In the case of a negative test instruction $-a$, the role of the value produced is reversed. In the case of a plain basic instruction a , the value produced is disregarded: execution always proceeds as if T is produced. The effect of a forward jump instruction $\#l$ is that execution proceeds with the l -th next instruction of the program concerned. If l equals 0 or the l -th next instruction does not exist, then $\#l$ results in deadlock. The effect of the termination instruction $!$ is that execution terminates.

PGA has the following constants and operators:

Table 5. Axioms of PGA

$(x ; y) ; z = x ; (y ; z)$	PGA1
$(x^n)^\omega = x^\omega$	PGA2
$x^\omega ; y = x^\omega$	PGA3
$(x ; y)^\omega = x ; (y ; x)^\omega$	PGA4

- for each $u \in \mathfrak{I}$, an *instruction* constant u ;
- the binary *concatenation* operator $_- ; _-$;
- the unary *repetition* operator $_-^\omega$.

Terms are built as usual. Throughout the paper, we assume that there are infinitely many variables, including x, y, z .

We use infix notation for concatenation and postfix notation for repetition.

Closed PGA terms are considered to denote programs. The intuition is that a program is in essence a non-empty, finite or periodic infinite sequence of primitive instructions.⁴ These sequences are called *single pass instruction sequences* because PGA has been designed to enable single pass execution of instruction sequences: each instruction can be dropped after it has been executed. Programs are considered to be equal if they represent the same single pass instruction sequence. The axioms for instruction sequence equivalence are given in Table 5. In this table, n stands for an arbitrary natural number greater than 0. For each $n > 0$, the term x^n is defined by induction on n as follows: $x^1 = x$ and $x^{n+1} = x ; x^n$. The *unfolding* equation $x^\omega = x ; x^\omega$ is derivable. Each closed PGA term is derivably equal to a term in *canonical form*, i.e. a term of the form P or $P ; Q^\omega$, where P and Q are closed PGA terms that do not contain the repetition operator.

Each closed PGA term is considered to denote a program of which the behaviour is a finite-state thread, taking the set \mathfrak{A} of basic instructions for the set \mathcal{A} of actions. The *thread extraction* operator $|_-$ assigns a thread to each program. The thread extraction operator is defined by the equations given in Table 6 (for $a \in \mathfrak{A}$, $l \in \mathbb{N}$ and $u \in \mathfrak{I}$) and the rule given in Table 7. This rule is expressed in terms of the *structural congruence* predicate $_- \cong _-$, which is defined by the formulas given in Table 8 (for $n, m, l \in \mathbb{N}$ and $u_1, \dots, u_n, v_1, \dots, v_{m+1} \in \mathfrak{I}$).

The equations given in Table 6 do not cover the case where there is a cyclic chain of forward jumps. Programs are structural congruent if they are the same after removing all chains of forward jumps in favour of single jumps. Because a cyclic chain of forward jumps corresponds to $\#0$, the rule from Table 7 can be read as follows: if x starts with a cyclic chain of forward jumps, then $|x|$ equals D . It is easy to see that the thread extraction operator assigns the same thread to structurally congruent programs. Therefore, the rule from Table 7 can be replaced by the following generalization: $x \cong y \Rightarrow |x| = |y|$.

⁴ A periodic infinite sequence is an infinite sequence with only finitely many subsequences.

Table 6. Defining equations for thread extraction operator

$ a = a \circ D$	$ \#l = D$
$ a; x = a \circ x $	$ \#0; x = D$
$ +a = a \circ D$	$ \#1; x = x $
$ +a; x = x \trianglelefteq a \triangleright \#2; x $	$ \#l+2; u = D$
$ -a = a \circ D$	$ \#l+2; u; x = \#l+1; x $
$ -a; x = \#2; x \trianglelefteq a \triangleright x $	$ \! = S$
	$ \! ; x = S$

Table 7. Rule for cyclic jump chains

$$\frac{x \cong \#0; y \Rightarrow |x| = D}{x \cong \#0; y \Rightarrow |x| = D}$$

Table 8. Defining formulas for structural congruence predicate

$\#n+1; u_1; \dots; u_n; \#0 \cong \#0; u_1; \dots; u_n; \#0$
$\#n+1; u_1; \dots; u_n; \#m \cong \#m+n+1; u_1; \dots; u_n; \#m$
$(\#n+l+1; u_1; \dots; u_n)^\omega \cong (\#l; u_1; \dots; u_n)^\omega$
$\#m+n+l+2; u_1; \dots; u_n; (v_1; \dots; v_{m+1})^\omega \cong$ $\#n+l+1; u_1; \dots; u_n; (v_1; \dots; v_{m+1})^\omega$
$x \cong x$
$x_1 \cong y_1 \wedge x_2 \cong y_2 \Rightarrow x_1; x_2 \cong y_1; y_2 \wedge x_1^\omega \cong y_1^\omega$

Let E be a finite guarded recursive specification over BTA, and let P_X be a closed PGA term for each $X \in V(E)$. Let E' be the set of equations that results from replacing in E all occurrences of X by $|P_X|$ for each $X \in V(E)$. If E' can be obtained by applications of axioms PGA1–PGA4, the defining equations for the thread extraction operator and the rule for cyclic jump chains, then $|P_X|$ is the solution of E for X . Such a finite guarded recursive specification can always be found. Thus, the behaviour of each closed PGA term, is a thread that is definable by a finite guarded recursive specification over BTA. Moreover, each finite guarded recursive specification over BTA can be translated to a closed PGA term of which the behaviour is the solution of the finite guarded recursive specification concerned (cf. Section 4 of [12]).

Closed PGA terms are loosely called PGA *programs*. PGA programs in which the repetition operator do not occur are called *finite* PGA programs.

4 The Program Notation PGLD

In this section, we review a program notation which is rooted in PGA. This program notation, called PGLD, belongs to a hierarchy of program notations

introduced in [3]. PGLD is close to existing assembly languages. It has absolute jump instructions and no explicit termination instruction.

In PGLD, like in PGA, it is assumed that there is a fixed but arbitrary finite set of *basic instructions* \mathfrak{A} . Again, the intuition is that the execution of a basic instruction a may modify a state and produces T or F at its completion.

PGLD has the following primitive instructions:

- for each $a \in \mathfrak{A}$, a *plain basic instruction* a ;
- for each $a \in \mathfrak{A}$, a *positive test instruction* $+a$;
- for each $a \in \mathfrak{A}$, a *negative test instruction* $-a$;
- for each $l \in \mathbb{N}$, a *direct absolute jump instruction* $\#\#l$.

PGLD programs have the form $u_1; \dots; u_k$, where u_1, \dots, u_k are primitive instructions of PGLD.

The plain basic instructions, the positive test instructions, and the negative test instructions are as in PGA. The effect of a direct absolute jump instruction $\#\#l$ is that execution proceeds with the l -th instruction of the program concerned. If $\#\#l$ is itself the l -th instruction, then deadlock occurs. If l equals 0 or l is greater than the length of the program, then termination occurs.

We define the meaning of PGLD programs by means of a function **pgld2pga** from the set of all PGLD programs to the set of all PGA programs. This function is defined by

$$\mathbf{pgld2pga}(u_1; \dots; u_k) = (\psi_1(u_1); \dots; \psi_k(u_k); !; !)^{\omega},$$

where the auxiliary functions ψ_j from the set of all primitive instructions of PGLD to the set of all primitive instructions of PGA are defined as follows ($1 \leq j \leq k$):

$$\begin{aligned} \psi_j(\#\#l) &= \#\#l - j && \text{if } j \leq l \leq k, \\ \psi_j(\#\#l) &= \#\#k + 2 - (j - l) && \text{if } 0 < l < j, \\ \psi_j(\#\#l) &= ! && \text{if } l = 0 \vee l > k, \\ \psi_j(u) &= u && \text{if } u \text{ is not a jump instruction.} \end{aligned}$$

Let P be a PGLD program. Then $\mathbf{pgld2pga}(P)$ represents the meaning of P as a PGA program. The intended behaviour of P is the behaviour of $\mathbf{pgld2pga}(P)$. That is, the *behaviour* of P , written $|P|_{\text{PGLD}}$, is $|\mathbf{pgld2pga}(P)|$.

We use the phrase *projection semantics* to refer to the approach to semantics followed in this section. The meaning function **pgld2pga** is called a *projection*.

5 Threads Interacting with Action Transforming Services

A thread may perform certain actions for having itself affected by some service or affecting some service. When processing an action performed by a thread, an action transforming service affects that thread in one of the following ways:

- (i) by returning a reply value to the thread at completion of the processing

of the action performed by the thread; (ii) by turning the processed action into another action. In this section, we introduce a mechanism that allows for services to affect threads in either way. The mechanism is a generalization of the use mechanism introduced in [8].⁵ We also introduce a complementary mechanism of that generalized use mechanism and an abstraction mechanism. The difference between the generalized use mechanism and the complementary mechanism is a matter of perspective: the former is concerned with the effect of services on threads and therefore produces threads, whereas the latter is concerned with the effect of threads on services and therefore produces services. The abstraction mechanism serves for concealment of the presence of internal actions, which arise primarily from the generalized use mechanism.

We will use the generalized use mechanism and the complementary mechanism of that mechanism for interactions between threads that are behaviours exhibited by sequential programs on execution and services that process actions that relate to molecular dynamics.

It is assumed that there is a fixed but arbitrary finite set of *foci* \mathcal{F} and a fixed but arbitrary finite set of *methods* \mathcal{M} . Each focus plays the role of a name of a service provided by the execution environment that can be requested to process a command. Each method plays the role of a command proper. For the set \mathcal{A} of actions, we take the set $\{f.m \mid f \in \mathcal{F}, m \in \mathcal{M}\}$. Performing an action $f.m$ is taken as making a request to the service named f to process command m .

We introduce yet another sort: the sort \mathbf{S} of *services*. \mathbf{S} is considered to stand for the set of all services. We identify services with pairs (H_1, H_2) , where $H_1: \mathcal{M}^+ \rightarrow \{\mathbf{T}, \mathbf{F}, \mathbf{M}, \mathbf{B}\}$ and $H_2: \mathcal{M}^+ \rightarrow \mathcal{A}_{\text{tau}}$, satisfying the following conditions:

$$\begin{aligned} & \forall m \in \mathcal{M} \bullet \\ & (\exists \alpha \in \mathcal{M}^* \bullet H_1(\alpha \circ \langle m \rangle) = \mathbf{M} \Rightarrow \forall \alpha' \in \mathcal{M}^* \bullet H_1(\alpha' \circ \langle m \rangle) \notin \{\mathbf{T}, \mathbf{F}\}) , \\ & \forall \alpha \in \mathcal{M}^+, m \in \mathcal{M} \bullet (H_1(\alpha) = \mathbf{B} \Rightarrow H_1(\alpha \circ \langle m \rangle) = \mathbf{B}) , \\ & \forall \alpha \in \mathcal{M}^+ \bullet (H_1(\alpha) \neq \mathbf{M} \Leftrightarrow H_2(\alpha) = \text{tau}) . \end{aligned}$$

Let H be a service, and let H_1 and H_2 be the unique functions such that $H = (H_1, H_2)$. Then we write $rf(H)$ and $af(H)$ for H_1 and H_2 , respectively. Given a service H and a method $m \in \mathcal{M}$, the *derived service* of H after processing m , written $\frac{\partial}{\partial m}H$, is defined by $rf(\frac{\partial}{\partial m}H)(\alpha) = rf(H)(\langle m \rangle \circ \alpha)$ and $af(\frac{\partial}{\partial m}H)(\alpha) = af(H)(\langle m \rangle \circ \alpha)$.

A service H can be understood as follows:

- if $rf(H)(\langle m \rangle) = \mathbf{T}$, then the request to process m is accepted by the service, a positive reply is produced, m is turned into tau , and the service proceeds as $\frac{\partial}{\partial m}H$;
- if $rf(H)(\langle m \rangle) = \mathbf{F}$, then the request to process m is accepted by the service, a negative reply is produced, m is turned into tau , and the service proceeds as $\frac{\partial}{\partial m}H$;

⁵ In later papers, the original use mechanism is also called thread-service composition.

- if $rf(H)(\langle m \rangle) = \mathbf{M}$, then the request to process m is accepted by the service, no reply is produced, m is turned into $af(H)(\langle m \rangle)$, and the service proceeds as $\frac{\partial}{\partial m}H$;
- if $rf(H)(\langle m \rangle) = \mathbf{B}$, then the request to process m is rejected by the service.

The three conditions imposed on services can be paraphrased as follows:

- if it is possible that no reply is produced at completion of the processing of a command, then it is impossible that a positive or negative reply is produced at completion of the processing of that command;
- after a request to process a command has been rejected, any request to process a command will be rejected;
- a reply is produced at completion of the processing of a command if and only if the command is turned into \mathbf{tau} .

We introduce the following additional constant and operators:

- the *divergent service* constant $\underline{\mathbf{D}} : \mathbf{S}$;
- for each $f \in \mathcal{F}$, the binary *use* operator $- /_f - : \mathbf{T} \times \mathbf{S} \rightarrow \mathbf{T}$;
- for each $f \in \mathcal{F}$, the binary *apply* operator $- \bullet_f - : \mathbf{T} \times \mathbf{S} \rightarrow \mathbf{S}$.

We use infix notation for use and apply.

$\underline{\mathbf{D}}$ is the unique service H with the property that $rf(H)(\alpha) = \mathbf{B}$ for all $\alpha \in \mathcal{M}^+$. The operators $- /_f -$ and $- \bullet_f -$ are complementary. Intuitively, $p /_f H$ is the thread that results from processing all actions performed by thread p that are of the form $f.m$ by the service H . When an action of the form $f.m$ performed by thread p is processed by the service H , that action is turned into another action and, if this action is \mathbf{tau} , postconditional composition is removed in favour of action prefixing on the basis of the reply value produced. Intuitively, $p \bullet_f H$ is the service that results from processing all basic actions performed by thread P that are of the form $f.m$ by service H . When an action of the form $f.m$ performed by thread p is processed by service H , that service is changed into $\frac{\partial}{\partial m}H$.

The axioms for the use and apply operators are given in Tables 9 and 10. In these tables, f and g stand for an arbitrary foci from \mathcal{F} , m stands for an arbitrary method from \mathcal{M} , and H is a variable of sort \mathbf{S} . Axioms TSU3 and TSU4 express that the action \mathbf{tau} and actions of the form $g.m$, where $f \neq g$, are not processed. Axioms TSU5–TSU7 express that a thread is affected by a service as described above when an action of the form $f.m$ performed by the thread is processed by the service. Axiom TSU8 expresses that deadlock takes place when an action to be processed is not accepted. Axiom TSU9 expresses that the divergent service does not accept any action. Axiom TSA3 expresses that a service is not affected by a thread when the action \mathbf{tau} is performed by the thread and axiom TSA4 expresses that a service is turned into the divergent service when an action of the form $g.m$, where $f \neq g$, is performed by the thread. Axioms TSA5–TSA7 express that a service is affected by a thread as described above when an action of the form $f.m$ performed by the thread is processed by the service. Axiom TSA8 expresses that a service is turned into the divergent

Table 9. Axioms for use operators

$S /_f H = S$	TSU1
$D /_f H = D$	TSU2
$\text{tau} \circ x /_f H = \text{tau} \circ (x /_f H)$	TSU3
$(x \sqsubseteq g.m \sqsupseteq y) /_f H = (x /_f H) \sqsubseteq g.m \sqsupseteq (y /_f H)$	if $f \neq g$ TSU4
$(x \sqsubseteq f.m \sqsupseteq y) /_f H = \text{tau} \circ (x /_f \frac{\partial}{\partial m} H)$	if $\text{rf}(H)(\langle m \rangle) = \text{T}$ TSU5
$(x \sqsubseteq f.m \sqsupseteq y) /_f H = \text{tau} \circ (y /_f \frac{\partial}{\partial m} H)$	if $\text{rf}(H)(\langle m \rangle) = \text{F}$ TSU6
$(x \sqsubseteq f.m \sqsupseteq y) /_f H = (x \sqsubseteq \text{af}(H)(\langle m \rangle) \sqsupseteq y) /_f \frac{\partial}{\partial m} H$	if $\text{rf}(H)(\langle m \rangle) = \text{M}$ TSU7
$(x \sqsubseteq f.m \sqsupseteq y) /_f H = D$	if $\text{rf}(H)(\langle m \rangle) = \text{B}$ TSU8
$(x \sqsubseteq f.m \sqsupseteq y) /_f \underline{D} = D$	TSU9

Table 10. Axioms for apply operators

$S \bullet_f H = H$	TSA1
$D \bullet_f H = \underline{D}$	TSA2
$(\text{tau} \circ x) \bullet_f H = x \bullet_f H$	TSA3
$(x \sqsubseteq g.m \sqsupseteq y) \bullet_f H = \underline{D}$	if $f \neq g$ TSA4
$(x \sqsubseteq f.m \sqsupseteq y) \bullet_f H = x \bullet_f \frac{\partial}{\partial m} H$	if $\text{rf}(H)(\langle m \rangle) = \text{T}$ TSA5
$(x \sqsubseteq f.m \sqsupseteq y) \bullet_f H = y \bullet_f \frac{\partial}{\partial m} H$	if $\text{rf}(H)(\langle m \rangle) = \text{F}$ TSA6
$(x \sqsubseteq f.m \sqsupseteq y) \bullet_f H = (x \sqsubseteq \text{af}(H)(\langle m \rangle) \sqsupseteq y) \bullet_f \frac{\partial}{\partial m} H$	if $\text{rf}(H)(\langle m \rangle) = \text{M}$ TSA7
$(x \sqsubseteq f.m \sqsupseteq y) \bullet_f H = \underline{D}$	if $\text{rf}(H)(\langle m \rangle) = \text{B}$ TSA8
$(x \sqsubseteq f.m \sqsupseteq y) \bullet_f \underline{D} = \underline{D}$	TSA9
$(\bigwedge_{n \geq 0} \pi_n(x) \bullet_f H = \underline{D}) \Rightarrow x \bullet_f H = \underline{D}$	TSA10

service when an action performed by the thread is not accepted. Axiom TSA9 expresses that the divergent service is not affected by a thread when an action of the form $f.m$ is performed by the thread.

We say that $p \bullet_f H$ is a *divergent thread application* if $\pi_n(p) \bullet_f H = \underline{D}$ for all $n \in \mathbb{N}$. Axiom TSA10 can be read as follows: if $p \bullet_f H$ is a divergent thread application, then it equals \underline{D} .

The use operators introduced in [8] deals in essence with services H where $\text{af}(H)(\alpha) = \text{tau}$ for all $\alpha \in \mathcal{M}^+$. For these services, the use operators introduced here coincide with those use operators.

Let T stand for either BTA, BTA+REC or BTA+REC+AIP. Then we will write $T + \text{TSI}$ for T , taking the set $\{f.m \mid f \in \mathcal{F}, m \in \mathcal{M}\}$ for \mathcal{A} , extended with the divergent service constant, the use and apply operators, and the axioms from Tables 9 and 10, with the exception of axiom TSA10 in the case where T does not stand for BTA+REC+AIP.

The action tau is an internal action whose presence matters. To conceal its presence in the case where it does not matter after all, we also introduce the unary *abstraction* operator $\tau_{\text{tau}} : \mathbf{T} \rightarrow \mathbf{T}$.

The axioms for the abstraction operator are given in Table 11. In this table,

Table 11. Axiom for abstraction

$\tau_{\mathbf{tau}}(\mathbf{S}) = \mathbf{S}$	TT1
$\tau_{\mathbf{tau}}(\mathbf{D}) = \mathbf{D}$	TT2
$\tau_{\mathbf{tau}}(x \trianglelefteq \mathbf{tau} \triangleright y) = \tau_{\mathbf{tau}}(x)$	TT3
$\tau_{\mathbf{tau}}(x \trianglelefteq a \triangleright y) = \tau_{\mathbf{tau}}(x) \trianglelefteq a \triangleright \tau_{\mathbf{tau}}(y)$	TT4

a stands for an arbitrary basic action from \mathcal{A} .

Abstraction can for instance be appropriate in the case where \mathbf{tau} arises from turning actions of an auxiliary nature into \mathbf{tau} on use of a service. An example of this case will occur in Section 9. Unlike the use mechanisms introduced in [8] and in the current paper, the use mechanism introduced in [9] incorporates abstraction.

Let T stand for either BTA, BTA+REC, BTA+REC+AIP, BTA+TSI, BTA+REC+TSI or BTA+REC+AIP+TSI. Then we will write T +ABSTR for T extended with the abstraction operator and the axioms from Table 11.

The equation $\tau_{\mathbf{tau}}(\mathbf{tau}^\omega) = \mathbf{D}$ is derivable from the axioms of BTA+REC+AIP+ABSTR.

To simplify matters, from now on the set $\{f.m \mid f \in \mathcal{F}, m \in \mathcal{M}\}$ is taken as the set \mathfrak{A} of basic instructions when PGA or PGLD is concerned. Thereby no real restriction is imposed on the set \mathfrak{A} : in the case where the cardinality of \mathcal{F} equals 1, all basic instructions have the same focus and the set \mathcal{M} of methods can be looked upon as the set \mathfrak{A} of basic instructions. We use the convention to omit foci from PGA programs in which all basic instructions have the same focus.

Strictly speaking, the propositions and theorems presented in this paper are proved in the algebraic theory obtained by: (i) combining PGA with BTA+REC+AIP+TSI+ABSTR, resulting in a theory with three sorts: a sort \mathbf{P} of programs, a sort \mathbf{T} of threads, and a sort \mathbf{S} of services; (ii) extending the result by taking $|-$ for an additional operator from sort \mathbf{P} to sort \mathbf{T} and taking the semantic equations and rule defining thread extraction for additional axioms.

6 State-Based Description of Services

In this section, we introduce the state-based approach to describe families of services that will be used in Section 7. This approach is similar to the approach to describe state machines introduced in [9].

In this approach, a family of services is described by

- a set of states S ;
- an effect function $eff : \mathcal{M} \times S \rightarrow S$;
- a yield function $yld : \mathcal{M} \times S \rightarrow \{\mathbf{T}, \mathbf{F}, \mathbf{M}, \mathbf{B}\}$;
- an action function $act : \mathcal{M} \times S \rightarrow \mathcal{A}_{\mathbf{tau}}$;

satisfying the following conditions:

$$\begin{aligned}
& \forall m \in \mathcal{M} \bullet (\exists s \in S \bullet yld(m, s) = \mathbf{M} \Rightarrow \forall s' \in S \bullet yld(m, s') \notin \{\mathbf{T}, \mathbf{F}\}) , \\
& \exists s \in S \bullet \forall m \in \mathcal{M} \bullet \\
& \quad (yld(m, s) = \mathbf{B} \wedge \forall s' \in S \bullet (yld(m, s') = \mathbf{B} \Rightarrow eff(m, s') = s)) , \\
& \forall m \in \mathcal{M}, s \in S \bullet (yld(m, s) \neq \mathbf{M} \Leftrightarrow act(m, s) = \mathbf{tau}) .
\end{aligned}$$

The set S contains the states in which the services may be, and the functions eff , yld and act give, for each method m and state s , the state, reply and action, respectively, that result from processing m in state s .

We define, for each $s \in S$, a cumulative effect function $ceff_s : \mathcal{M}^* \rightarrow S$ in terms of s and eff as follows:

$$\begin{aligned}
ceff_s(\langle \rangle) &= s , \\
ceff_s(\alpha \curvearrowright \langle m \rangle) &= eff(m, ceff_s(\alpha)) .
\end{aligned}$$

We define, for each $s \in S$, a service H_s in terms of $ceff_s$, yld and act as follows:

$$\begin{aligned}
rf(H_s)(\alpha \curvearrowright \langle m \rangle) &= yld(m, ceff_s(\alpha)) , \\
af(H_s)(\alpha \curvearrowright \langle m \rangle) &= act(m, ceff_s(\alpha)) .
\end{aligned}$$

H_s is called the service with *initial state* s described by S , eff , yld and act . We say that $\{H_s \mid s \in S\}$ is the *family of services* described by S , eff , yld and act .

The conditions that are imposed on S , eff , yld and act imply that, for each $s \in S$, H_s is a service indeed. It is worth mentioning that $\frac{\partial}{\partial m} H_s = H_{eff(m, s)}$, $rf(H_s)(\langle m \rangle) = yld(m, s)$, and $af(H_s)(\langle m \rangle) = act(m, s)$.

7 Services for Molecular Dynamics

In this section, we describe a family of services which concerns molecular dynamics. The formal description given here elaborates on an informal description of molecular dynamics given in [1].

The states of molecular dynamics services resemble collections of molecules composed of atoms and the methods of molecular dynamics services transform the structure of molecules like in chemical reactions. An atom can have *fields* and each of those fields can contain an atom. An atom together with the ones it has links to via fields can be viewed as a submolecule, and a submolecule that is not contained in a larger submolecule can be viewed as a molecule. Thus, the collection of molecules that make up a state can be viewed as a fluid. By means of methods, new atoms can be created, fields can be added to and removed from atoms, and the contents of fields of atoms can be examined and modified. A few methods use a *spot* to put an atom in or to get an atom from. By means of methods, the contents of spots can be compared and modified as well. Creating an atom is thought of as turning an element of a given set of *proto-atoms* into an atom. If there are no proto-atoms left, then atoms can no longer be created.

It is assumed that a finite set **Spot** of *spots* such that $\mathcal{F}, \mathcal{M} \subseteq \mathbf{Spot}$, a total order \leq on **Spot**, and a finite set **Field** of *fields* have been given. It is further

assumed that a countable set PAtom of *proto-atoms* such that $\perp \notin \text{PAtom}$ and a bijection $\text{patom} : [1, \text{card}(\text{PAtom})] \rightarrow \text{PAtom}$ have been given. Although the set of proto-atoms may be infinite, there exists at any time only a finite number of atoms. Each of those atoms has only a finite number of fields. Molecular dynamics services accept the following methods:

- for each $s \in \text{Spot}$, a *create atom method* $s!$;
- for each $s, s' \in \text{Spot}$, a *set spot method* $s = s'$;
- for each $s \in \text{Spot}$, a *clear spot method* $s = 0$;
- for each $s, s' \in \text{Spot}$, an *equality test method* $s == s'$;
- for each $s \in \text{Spot}$, an *undefinedness test method* $s == 0$;
- for each $s \in \text{Spot}$ and $v \in \text{Field}$, a *add field method* s/v ;
- for each $s \in \text{Spot}$ and $v \in \text{Field}$, a *remove field method* $s \backslash v$;
- for each $s \in \text{Spot}$ and $v \in \text{Field}$, a *has field method* $s|v$;
- for each $s \in \text{Spot}$ and $v \in \text{Field}$, a *set field method* $s.v = s'$;
- for each $s \in \text{Spot}$ and $v \in \text{Field}$, a *get field method* $s = s'.v$;
- for each $s, s' \in \text{Spot}$, a *generate action method* $\%s.\%s'$.

We write \mathcal{M}_{gen} for the set of all generate action methods and \mathcal{M}_{md} for the set of all methods of molecular dynamics services. It is assumed that $\mathcal{M}_{\text{md}} \subseteq \mathcal{M}$.

The states of molecular dynamics services comprise the contents of all spots, the fields of the existing atoms, and the contents of those fields. The methods accepted by molecular dynamics services can be explained as follows:

- $s!$: if an atom can be created, then the contents of spot s becomes a newly created atom and the reply is T; otherwise, nothing changes and the reply is F;
- $s = s'$: the contents of spot s becomes the same as the contents of spot s' and the reply is T;
- $s = 0$: the contents of spot s becomes undefined and the reply is T;
- $s == s'$: if the contents of spot s equals the contents of spot s' , then nothing changes and the reply is T; otherwise, nothing changes and the reply is F;
- $s == 0$: if the contents of spot s is undefined, then nothing changes and the reply is T; otherwise, nothing changes and the reply is F;
- s/v : if the contents of spot s is an atom and v is not yet a field of that atom, then v is added (with undefined contents) to the fields of that atom and the reply is T; otherwise, nothing changes and the reply is F;
- $s \backslash v$: if the contents of spot s is an atom and v is a field of that atom, then v is removed from the fields of that atom and the reply is T; otherwise, nothing changes and the reply is F;
- $s|v$: if the contents of spot s is an atom and v is a field of that atom, then nothing changes and the reply is T; otherwise, nothing changes and the reply is F;
- $s.v = s'$: if the contents of spot s is an atom and v is a field of that atom, then the contents of that field becomes the same as the contents of spot s' and the reply is T; otherwise, nothing changes and the reply is F;

- $s = s'.v$: if the contents of spot s' is an atom and v is a field of that atom, then the contents of spot s becomes the same as the contents of that field and the reply is T ; otherwise, nothing changes and the reply is F ;
- $\%s.\%s'$: if the contents of spots s and s' are atoms and there exist $f' \in \mathcal{F}$ and $m' \in \mathcal{M}$ such that the contents of spot f' equals the contents of spot s and the contents of spot m' equals the contents of spot s' , then nothing changes, there is no reply, and $\%s.\%s'$ is turned into $f.m$ where f and m are the least $f' \in \mathcal{F}$ and $m' \in \mathcal{M}$ with respect to \leq that satisfy the conditions formulated above; otherwise, $\%s.\%s'$ is rejected.

In the explanation given above, wherever we say that the reply is T or the reply is F , this is meant to imply that the method concerned is turned into tau . Moreover, wherever we say that the contents of a spot or field becomes the same as the contents of another spot or field, this is meant to imply that the former contents becomes undefined if the latter contents is undefined.

Let

$$\begin{aligned}
SS &= \text{Spot} \rightarrow (\text{PAtom} \cup \{\perp\}) , \\
AS &= \bigcup_{A \in \mathcal{P}_{\text{fin}}(\text{PAtom})} (A \rightarrow \bigcup_{F \in \mathcal{P}_{\text{fin}}(\text{Field})} (F \rightarrow (\text{PAtom} \cup \{\perp\}))) , \\
MDS &= \{(\sigma, \alpha) \in SS \times AS \mid \text{rng}(\sigma) \subseteq \text{dom}(\alpha) \cup \{\perp\} \wedge \\
&\quad \forall a \in \text{dom}(\alpha) \bullet \text{rng}(\alpha(a)) \subseteq \text{dom}(\alpha) \cup \{\perp\}\} , \\
s &\in MDS .
\end{aligned}$$

Then we write \mathcal{MD}_s for the service with initial state s described by $S = MDS \cup \{\uparrow\}$, where $\uparrow \notin MDS$, and the functions eff , yld and act defined in Tables 12, 13 and 14.⁶ In these tables, the functions $\mathit{new} : \mathcal{P}_{\text{fin}}(\text{PAtom}) \rightarrow (\text{PAtom} \cup \{\perp\})$, $\mathit{gencnd} : SS \times \text{Spot} \times \text{Spot} \rightarrow \{\mathsf{T}, \mathsf{F}\}$, and $\mathit{genact} : SS \times \text{Spot} \times \text{Spot} \rightarrow \mathcal{A}$ are used. These functions are defined as follows:

$$\begin{aligned}
\mathit{new}(A) &= \mathit{patom}(m+1) \text{ if } m < \text{card}(\text{PAtom}) , \\
\mathit{new}(A) &= \perp \quad \quad \quad \text{if } m \geq \text{card}(\text{PAtom}) ,
\end{aligned}$$

where $m = \max\{n \mid \mathit{patom}(n) \in A\}$;

$$\begin{aligned}
\mathit{gencnd}(\sigma, s, s') &= \mathsf{T} \text{ iff} \\
&\sigma(s) \neq \perp \wedge \sigma(s') \neq \perp \wedge \exists f \in \mathcal{F} \bullet \sigma(f) = \sigma(s) \wedge \exists m \in \mathcal{M} \bullet \sigma(m) = \sigma(s') ;
\end{aligned}$$

$$\mathit{genact}(\sigma, s, s') = f.m ,$$

⁶ We use the following notation for functions: $[]$ for the empty function; $[d \mapsto r]$ for the function f with $\text{dom}(f) = \{d\}$ such that $f(d) = r$; $f \oplus g$ for the function h with $\text{dom}(h) = \text{dom}(f) \cup \text{dom}(g)$ such that for all $d \in \text{dom}(h)$, $h(d) = f(d)$ if $d \notin \text{dom}(g)$ and $h(d) = g(d)$ otherwise; and $f \triangleleft D$ for the function g with $\text{dom}(g) = \text{dom}(f) \setminus D$ such that for all $d \in \text{dom}(g)$, $g(d) = f(d)$.

Table 12. Effect function for molecular dynamics services

$eff(s!, (\sigma, \alpha)) =$	
$(\sigma \oplus [s \mapsto new(\text{dom}(\alpha))], \alpha \oplus [new(\text{dom}(\alpha)) \mapsto []])$	if $new(\text{dom}(\alpha)) \neq \perp$
$eff(s!, (\sigma, \alpha)) = (\sigma, \alpha)$	if $new(\text{dom}(\alpha)) = \perp$
$eff(s = s', (\sigma, \alpha)) = (\sigma \oplus [s \mapsto \sigma(s')], \alpha)$	
$eff(s = 0, (\sigma, \alpha)) = (\sigma \oplus [s \mapsto \perp], \alpha)$	
$eff(s == s', (\sigma, \alpha)) = (\sigma, \alpha)$	
$eff(s == 0, (\sigma, \alpha)) = (\sigma, \alpha)$	
$eff(s/v, (\sigma, \alpha)) =$	
$(\sigma, \alpha \oplus [\sigma(s) \mapsto \alpha(\sigma(s)) \oplus [v \mapsto \perp]])$	if $\sigma(s) \neq \perp \wedge v \notin \text{dom}(\alpha(\sigma(s)))$
$eff(s/v, (\sigma, \alpha)) = (\sigma, \alpha)$	if $\sigma(s) = \perp \vee v \in \text{dom}(\alpha(\sigma(s)))$
$eff(s \setminus v, (\sigma, \alpha)) = (\sigma, \alpha \oplus [\sigma(s) \mapsto \alpha(\sigma(s)) \triangleleft \{v\}])$	if $\sigma(s) \neq \perp \wedge v \in \text{dom}(\alpha(\sigma(s)))$
$eff(s \setminus v, (\sigma, \alpha)) = (\sigma, \alpha)$	if $\sigma(s) = \perp \vee v \notin \text{dom}(\alpha(\sigma(s)))$
$eff(s v, (\sigma, \alpha)) = (\sigma, \alpha)$	
$eff(s.v = s', (\sigma, \alpha)) =$	
$(\sigma, \alpha \oplus [\sigma(s) \mapsto \alpha(\sigma(s)) \oplus [v \mapsto \sigma(s')]])$	if $\sigma(s) \neq \perp \wedge v \in \text{dom}(\alpha(\sigma(s)))$
$eff(s.v = s', (\sigma, \alpha)) = (\sigma, \alpha)$	if $\sigma(s) = \perp \vee v \notin \text{dom}(\alpha(\sigma(s)))$
$eff(s = s'.v, (\sigma, \alpha)) = (\sigma \oplus [s \mapsto \alpha(\sigma(s'))(v)], \alpha)$	if $\sigma(s') \neq \perp \wedge v \in \text{dom}(\alpha(\sigma(s')))$
$eff(s = s'.v, (\sigma, \alpha)) = (\sigma, \alpha)$	if $\sigma(s') = \perp \vee v \notin \text{dom}(\alpha(\sigma(s')))$
$eff(\%s.\%s', (\sigma, \alpha)) = (\sigma, \alpha)$	if $gencnd(\sigma, s, s') = \text{T}$
$eff(\%s.\%s', (\sigma, \alpha)) = \uparrow$	if $gencnd(\sigma, s, s') = \text{F}$
$eff(m, (\sigma, \alpha)) = \uparrow$	if $m \notin \mathcal{M}_{\text{md}}$
$eff(m, \uparrow) = \uparrow$	

where $f = \min\{f' \in \mathcal{F} \mid \sigma(f') = \sigma(s)\}$ and $m = \min\{m' \in \mathcal{M} \mid \sigma(m') = \sigma(s)\}$. We write \mathcal{MD} for the family of services described by S , eff , yld , and act . We write $\mathcal{MD}_{\text{init}}$ for the service from this family of which the initial state is the unique $(\sigma, \alpha) \in S$ such that $\text{dom}(\alpha) = []$.

Let $(\sigma, \alpha) \in S$, let $s \in \text{Spot}$, let $a \in \text{dom}(\alpha)$, and let $v \in \text{dom}(\alpha(a))$. Then $\sigma(s)$ is the contents of spot s if $\sigma(s) \neq \perp$, v is a field of atom a , and $\alpha(a)(v)$ is the contents of field v of atom a if $\alpha(a)(v) \neq \perp$. The contents of spot s is undefined if $\sigma(s) = \perp$, and the contents of field v of atom a is undefined if $\alpha(a)(v) = \perp$. Notice that $\text{dom}(\alpha)$ is taken for the set of all existing atoms. Therefore, the contents of each spot, i.e. each element of $\text{rng}(\sigma)$, must be in $\text{dom}(\alpha)$ if the contents is defined. Moreover, for each existing atom a , the contents of each of its fields, i.e. each element of $\text{rng}(\alpha(a))$, must be in $\text{dom}(\alpha)$ if the contents is defined. Molecular dynamics services get into state \uparrow when refusing a request to process a command. The molecular dynamics service with initial state \uparrow is the same as the divergent service $\underline{\mathbf{D}}$. The function new turns proto-atoms into atoms. After all proto-atoms have been turned into atoms, new yields \perp . This can only

Table 13. Yield function for molecular dynamics services

$yld(s!, (\sigma, \alpha)) = \mathbf{T}$	if $new(\text{dom}(\alpha)) \neq \perp$
$yld(s!, (\sigma, \alpha)) = \mathbf{F}$	if $new(\text{dom}(\alpha)) = \perp$
$yld(s = s', (\sigma, \alpha)) = \mathbf{T}$	
$yld(s = 0, (\sigma, \alpha)) = \mathbf{T}$	
$yld(s == s', (\sigma, \alpha)) = \mathbf{T}$	if $\sigma(s) = \sigma(s')$
$yld(s == s', (\sigma, \alpha)) = \mathbf{F}$	if $\sigma(s) \neq \sigma(s')$
$yld(s == 0, (\sigma, \alpha)) = \mathbf{T}$	if $\sigma(s) = \perp$
$yld(s == 0, (\sigma, \alpha)) = \mathbf{F}$	if $\sigma(s) \neq \perp$
$yld(s/v, (\sigma, \alpha)) = \mathbf{T}$	if $\sigma(s) \neq \perp \wedge v \notin \text{dom}(\alpha(\sigma(s)))$
$yld(s/v, (\sigma, \alpha)) = \mathbf{F}$	if $\sigma(s) = \perp \vee v \in \text{dom}(\alpha(\sigma(s)))$
$yld(s \setminus v, (\sigma, \alpha)) = \mathbf{T}$	if $\sigma(s) \neq \perp \wedge v \in \text{dom}(\alpha(\sigma(s)))$
$yld(s \setminus v, (\sigma, \alpha)) = \mathbf{F}$	if $\sigma(s) = \perp \vee v \notin \text{dom}(\alpha(\sigma(s)))$
$yld(s v, (\sigma, \alpha)) = \mathbf{T}$	if $\sigma(s) \neq \perp \wedge v \in \text{dom}(\alpha(\sigma(s)))$
$yld(s v, (\sigma, \alpha)) = \mathbf{F}$	if $\sigma(s) = \perp \vee v \notin \text{dom}(\alpha(\sigma(s)))$
$yld(s.v = s', (\sigma, \alpha)) = \mathbf{T}$	if $\sigma(s) \neq \perp \wedge v \in \text{dom}(\alpha(\sigma(s)))$
$yld(s.v = s', (\sigma, \alpha)) = \mathbf{F}$	if $\sigma(s) = \perp \vee v \notin \text{dom}(\alpha(\sigma(s)))$
$yld(s = s'.v, (\sigma, \alpha)) = \mathbf{T}$	if $\sigma(s') \neq \perp \wedge v \in \text{dom}(\alpha(\sigma(s')))$
$yld(s = s'.v, (\sigma, \alpha)) = \mathbf{F}$	if $\sigma(s') = \perp \vee v \notin \text{dom}(\alpha(\sigma(s')))$
$yld(\%s.\%s', (\sigma, \alpha)) = \mathbf{M}$	if $genrnd(\sigma, s, s') = \mathbf{T}$
$yld(\%s.\%s', (\sigma, \alpha)) = \mathbf{B}$	if $genrnd(\sigma, s, s') = \mathbf{F}$
$yld(m, (\sigma, \alpha)) = \mathbf{B}$	if $m \notin \mathcal{M}_{\text{md}}$
$yld(m, \uparrow) = \mathbf{B}$	

Table 14. Action function for molecular dynamics services

$act(m, (\sigma, \alpha)) = \mathbf{tau}$	if $m \notin \mathcal{M}_{\text{gen}}$
$act(\%s.\%s', (\sigma, \alpha)) = genact(\sigma, s, s')$	if $genrnd(\sigma, s, s') = \mathbf{T}$
$act(\%s.\%s', (\sigma, \alpha)) = \mathbf{tau}$	if $genrnd(\sigma, s, s') = \mathbf{F}$
$act(m, \uparrow) = \mathbf{tau}$	

happen if the number of proto-atoms is finite. The initial state of $\mathcal{MD}_{\text{init}}$ is the unique state in which no proto-atoms have been turned into atoms yet.

8 Representing Programs by Molecules

In this section, we associate each PGLD program with a PGA program for constructing a representation of the PGLD program by a molecule.

Let $u_1 ; \dots ; u_k$ be a PGLD program in which the foci f_1, \dots, f_n and the methods $m_1, \dots, m_{n'}$ occur. The idea is that:

- an atom is created for each of the foci f_1, \dots, f_n and the methods $m_1, \dots, m_{n'}$, using the focus or method itself as the spot into which the atom concerned is brought on its creation, and an atom is created for each of the instructions u_1, \dots, u_k ;
- each atom that corresponds to an instruction is linked via fields to other atoms as follows:
 - if the corresponding instruction is of the form $f.m$, $+f.m$ or $-f.m$, then the atom is linked to the atoms that correspond to the focus and method concerned and the atoms that correspond to the instructions with which execution must proceed in the cases of a positive and a negative reply;
 - if the corresponding instruction is of the form $\#\#l$ and $1 \leq l \leq k$, then the atom is linked to the atom that corresponds to the instruction with which execution must proceed, i.e. the l -th instruction;
- to prevent that an exception must be made of the instruction u_k in the case where it is of the form $f.m$, $+f.m$ or $-f.m$, two additional atoms are created that are not linked to other atoms;
- the atom that corresponds to the first instruction is brought into spot s .

Notice that, if an atom corresponds to an instruction of the form $\#\#l$ and not $1 \leq l \leq k$, then the atom is not linked to another atom.

We make the assumptions that \mathcal{F} and \mathcal{M} are disjoint and that PAtom is infinite. Under these assumptions, atom creation always leads to a positive reply and no test instructions are needed in the programs for constructing representations of PGLD programs. The first assumption is only made because it permits this simplification. The second assumption is primarily made because there will be PGLD programs for which there are no PGA programs for constructing their representation if PAtom is finite.

We define the PGA programs for constructing representations of PGLD programs by means of a function pgld2md from the set of all PGLD programs to the set of all PGA programs. This function is defined by

$$\begin{aligned}
 \text{pgld2md}(u_1; \dots; u_k) &= f_1!; \dots; f_n!; m_1!; \dots; m_{n'}!; s_1!; \dots; s_{k+2}!; \\
 &\quad \rho_1(u_1); \dots; \rho_k(u_k); \\
 &\quad s_{k+1}/\text{stop}; s_{k+2}/\text{stop}; \\
 &\quad s = s_1; !,
 \end{aligned}$$

where the auxiliary functions ρ_j from the set of all primitive instructions of PGLD to the set of all PGA programs are defined as follows ($1 \leq j \leq k$):

$$\begin{aligned}
 \rho_j(f.m) &= s_j/\text{focus}; s_j/\text{meth}; s_j/\text{pos}; s_j/\text{neg}; \\
 &\quad s_j.\text{focus} = f; s_j.\text{meth} = m; s_j.\text{pos} = s_{j+1}; s_j.\text{neg} = s_{j+1}, \\
 \rho_j(+f.m) &= s_j/\text{focus}; s_j/\text{meth}; s_j/\text{pos}; s_j/\text{neg}; \\
 &\quad s_j.\text{focus} = f; s_j.\text{meth} = m; s_j.\text{pos} = s_{j+1}; s_j.\text{neg} = s_{j+2}, \\
 \rho_j(-f.m) &= s_j/\text{focus}; s_j/\text{meth}; s_j/\text{pos}; s_j/\text{neg}; \\
 &\quad s_j.\text{focus} = f; s_j.\text{meth} = m; s_j.\text{pos} = s_{j+2}; s_j.\text{neg} = s_{j+1},
 \end{aligned}$$

$$\begin{aligned}\rho_j(\#\#l) &= s_j/jmp ; s_j.jmp = s_l & \text{if } 1 \leq l \leq k, \\ \rho_j(\#\#l) &= s_j/stop & \text{if } \neg(1 \leq l \leq k)\end{aligned}$$

and

- $f_1, \dots, f_n \in \mathcal{F}$ are the different foci that occur in $u_1 ; \dots ; u_k$;
- $m_1, \dots, m_{n'} \in \mathcal{M}$ are the different methods that occur in $u_1 ; \dots ; u_k$;
- $s, s_1, \dots, s_{k+2} \in \text{Spot} \setminus (\mathcal{F} \cup \mathcal{M})$.

In this definition, the omitted focus is considered to be md .

The properties stated in the following proposition show that it is easy to retrieve the PGLD program P from its representation constructed by $\text{pgld2md}(P)$.

Proposition 1. *Let $P = u_1 ; \dots ; u_k$ be a PGLD program, and let (σ, α) be such that $\mathcal{MD}_{(\sigma, \alpha)} = |\text{pgld2md}(P)| \bullet_{\text{md}} \mathcal{MD}_{\text{init}}$. Then $\text{yld}(s == s_1, (\sigma, \alpha)) = \top$ and for all $j, l \in [1, k]$, $f \in \mathcal{F}$ and $m \in \mathcal{M}$:*

- $u_j = f.m$ iff
 - $\text{yld}(u = s_j.\text{focus}, (\sigma, \alpha)) = \top$, $\text{yld}(u == f, \text{eff}(u = s_j.\text{focus}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(v = s_j.\text{meth}, (\sigma, \alpha)) = \top$, $\text{yld}(v == m, \text{eff}(u = s_j.\text{meth}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(s = s_j.\text{pos}, (\sigma, \alpha)) = \top$, $\text{yld}(s == s_{j+1}, \text{eff}(s = s_j.\text{pos}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(s = s_j.\text{neg}, (\sigma, \alpha)) = \top$, $\text{yld}(s == s_{j+1}, \text{eff}(s = s_j.\text{neg}, (\sigma, \alpha))) = \top$;
- $u_j = +f.m$ iff
 - $\text{yld}(u = s_j.\text{focus}, (\sigma, \alpha)) = \top$, $\text{yld}(u == f, \text{eff}(u = s_j.\text{focus}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(v = s_j.\text{meth}, (\sigma, \alpha)) = \top$, $\text{yld}(v == m, \text{eff}(u = s_j.\text{meth}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(s = s_j.\text{pos}, (\sigma, \alpha)) = \top$, $\text{yld}(s == s_{j+1}, \text{eff}(s = s_j.\text{pos}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(s = s_j.\text{neg}, (\sigma, \alpha)) = \top$, $\text{yld}(s == s_{j+2}, \text{eff}(s = s_j.\text{neg}, (\sigma, \alpha))) = \top$;
- $u_j = -f.m$ iff
 - $\text{yld}(u = s_j.\text{focus}, (\sigma, \alpha)) = \top$, $\text{yld}(u == f, \text{eff}(u = s_j.\text{focus}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(v = s_j.\text{meth}, (\sigma, \alpha)) = \top$, $\text{yld}(v == m, \text{eff}(u = s_j.\text{meth}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(s = s_j.\text{pos}, (\sigma, \alpha)) = \top$, $\text{yld}(s == s_{j+2}, \text{eff}(s = s_j.\text{pos}, (\sigma, \alpha))) = \top$,
 - $\text{yld}(s = s_j.\text{neg}, (\sigma, \alpha)) = \top$, $\text{yld}(s == s_{j+1}, \text{eff}(s = s_j.\text{neg}, (\sigma, \alpha))) = \top$;
- $u_j = \#\#l$ iff
 - $\text{yld}(s = s_j.\text{jmp}, (\sigma, \alpha)) = \top$, $\text{yld}(s == s_l, \text{eff}(s = s_j.\text{jmp}, (\sigma, \alpha))) = \top$;
- $u_j = \#\#l'$ for some $l' \notin [1, k]$ iff $\text{yld}(s_j|stop, (\sigma, \alpha)) = \top$.

Proof. From the assumption that \mathcal{F} and \mathcal{M} are disjoint, the assumption that PAtom is infinite, and the definition of pgld2md , it follows easily that:

- all atom creations are successful;
- the content of each spot used in an atom creation does not change after the atom creation concerned;
- $\rho_j(u_j)$ modifies the atom that is the contents of spot s_j only ($1 \leq j \leq k$), $s_{k+1}/stop$ modifies the atom that is the contents of spot s_{k+1} only, and $s_{k+2}/stop$ modifies the atom that is the contents of spot s_{k+2} only.

From this, the properties stated in the proposition follow straightforwardly by case distinction as in the definition of ρ_j . \square

9 Interpreting Programs Represented by Molecules

In this section, we introduce a PGA program for interpreting PGLD programs represented by molecules.

The idea is that:

- if the atom to be interpreted has the field *stop*, then the interpretation terminates;
- if the atom to be interpreted has the field *jmp*, then the interpretation proceeds with the atom to which it is linked via this field;
- otherwise, first the basic instruction represented by the atoms to which it is linked via the fields *focus* and *meth* is executed and then the interpretation proceeds with the atom to which it is linked via the field *pos* or the field *neg*, depending upon the reply being positive or negative.

The following is the PGA program for interpreting PGLD programs represented by molecules:

```
(+s|stop;!;
+s|jmp;#9;
u = s.focus; v = s.meth; +%u.%v; #3; s = s.neg; #4; s = s.pos; #2;
s = s.jmp)ω,
```

where $u, v \in \mathbf{Spot} \setminus (\mathcal{F} \cup \mathcal{M})$. Again, the omitted focus is considered to be **md**. Below, we write I for this PGA program.

The program I interprets PGLD programs correctly in the sense that, for all PGLD programs P , the intended behaviour of P under execution coincides with the behaviour of the interpreter under execution on interaction with the molecular dynamics service that holds the representation of P constructed by the program $\mathbf{pgld2md}(P)$ when abstracted from **tau**. This is stated rigorously by Theorem 1 below. The theorem is preceded by Proposition 2 which is used in the proof of the theorem. The proposition concerns the local use of the spots u and v in I .

For convenience, we introduce some abbreviations. Let $P = u_1; \dots; u_k$ be a PGLD program and let $i \in [1, k]$. Then we write $\mathcal{MD}_{\text{repr}(P)}$ for $|\mathbf{pgld2md}(P)| \bullet_{\mathbf{md}} \mathcal{MD}_{\text{init}}$ and $\mathcal{MD}_{\text{repr}(i,P)}$ for $|s = s_i; !| \bullet_{\mathbf{md}} \mathcal{MD}_{\text{repr}(P)}$.

Proposition 2. *Let $P = u_1; \dots; u_k$ be a PGLD program, let $i \in [1, k]$, and let $\mathcal{MD}'_{\text{repr}(i,P)}$ be such that $\mathcal{MD}_{\text{repr}(i,P)} = |u = 0; v = 0; !| \bullet_{\mathbf{md}} \mathcal{MD}'_{\text{repr}(i,P)}$. Then $\tau_{\text{tau}}(|I| /_{\mathbf{md}} \mathcal{MD}_{\text{repr}(i,P)}) = \tau_{\text{tau}}(|I| /_{\mathbf{md}} \mathcal{MD}'_{\text{repr}(i,P)})$.*

Proof. By AIP, it is sufficient to prove that for all $n \in \mathbb{N}$, for all $i \in [1, k]$, $\pi_n(\tau_{\text{tau}}(|I| /_{\mathbf{md}} \mathcal{MD}_{\text{repr}(i,P)})) = \pi_n(\tau_{\text{tau}}(|I| /_{\mathbf{md}} \mathcal{MD}'_{\text{repr}(i,P)}))$. This property is straightforward to prove by induction on n . \square

Theorem 1. *For all PGLD programs P :*

$$|P|_{\text{PGLD}} = \tau_{\text{tau}}(|I| /_{\mathbf{md}} \mathcal{MD}_{\text{repr}(P)}) .$$

Proof. In the proof, we make use of an auxiliary function $|_, _|$ from the set of all natural numbers and the set of all PGLD programs to the set of all closed terms of sort **T**. It gives, for each natural number i and PGLD program $u_1 ; \dots ; u_k$, a closed term of sort **T** that denotes the behaviour of $u_1 ; \dots ; u_k$ when executed from position i if $1 \leq i \leq k$ and **S** otherwise. This function is defined as follows:

$$\begin{aligned} |i, u_1 ; \dots ; u_k| &= |\psi_i(u_i) ; \dots ; \psi_k(u_k) ; ! ; ! ; (\psi_1(u_1) ; \dots ; \psi_k(u_k) ; ! ; !)^\omega| && \text{if } 1 \leq i \leq k, \\ |i, u_1 ; \dots ; u_k| &= \mathbf{S} && \text{if } \neg 1 \leq i \leq k \end{aligned}$$

(where ψ_j is as in the definition of `pgld2pga`). It follows easily from the definition of $|_, _|$ and the axioms of PGA that if $1 \leq i \leq k$:

$$\begin{aligned} |i, u_1 ; \dots ; u_k| &= a \circ |i + 1, u_1 ; \dots ; u_k| && \text{if } u_i = a, \\ |i, u_1 ; \dots ; u_k| &= |i + 1, u_1 ; \dots ; u_k| \trianglelefteq a \trianglerighteq |i + 2, u_1 ; \dots ; u_k| && \text{if } u_i = +a, \\ |i, u_1 ; \dots ; u_k| &= |i + 2, u_1 ; \dots ; u_k| \trianglelefteq a \trianglerighteq |i + 1, u_1 ; \dots ; u_k| && \text{if } u_i = -a, \\ |i, u_1 ; \dots ; u_k| &= |l, u_1 ; \dots ; u_k| && \text{if } u_i = \#\#l. \end{aligned}$$

Let $P = u_1 ; \dots ; u_k$ be a PGLD program, let

$$\begin{aligned} T &= \{|i, P| \mid i \in [1, k]\}, \\ T' &= \{\tau_{\text{tau}}(|I| /_{\text{md}} \mathcal{MD}_{\text{repr}(i, P)}) \mid i \in [1, k]\}, \end{aligned}$$

and let $\beta : T \rightarrow T'$ be the bijection defined by

$$\beta(|i, P|) = \tau_{\text{tau}}(|I| /_{\text{md}} \mathcal{MD}_{\text{repr}(i, P)}).$$

For each closed term p' of sort **T**, write $\beta^*(p')$ for p' with, for all $p \in T$, all occurrences of p in p' replaced by $\beta(p)$. Then, using the equations concerning the auxiliary function $|_, _|$ given above and Propositions 1 and 2, it is straightforward to prove that there exists a set E consisting of one derivable equation $p = p'$ for each $p \in T$ such that, for all equations $p = p'$ in E :

- the equation $\beta(p) = \beta^*(p')$ is derivable;
- $p' \in T$ only if p' can be rewritten to a $p'' \notin T$ using the equations in E from left to right.

Because $|P|_{\text{PGLD}} = |1, P|$ and $\tau_{\text{tau}}(|I| /_{\text{md}} \mathcal{MD}_{\text{repr}(P)}) = \tau_{\text{tau}}(|I| /_{\text{md}} \mathcal{MD}_{\text{repr}(1, P)})$, this means that $|P|_{\text{PGLD}}$ and $\tau_{\text{tau}}(|I| /_{\text{md}} \mathcal{MD}_{\text{repr}(P)})$ are solutions of the same guarded recursive specification. Because guarded recursive specifications have unique solutions, it follows immediately that $|P|_{\text{PGLD}} = \tau_{\text{tau}}(|I| /_{\text{md}} \mathcal{MD}_{\text{repr}(P)})$. \square

In the program `pgld2md`($u_1 ; \dots ; u_k$), we could have replaced $s = s_1 ; !$ by $s = s_1 ; s_1 = 0 ; \dots ; s_{k+2} = 0 ; !$ because the program I does not use the spots s_1, \dots, s_{k+2} . However, we cannot conceive a proof of Theorem 1 in which these auxiliary spots are not found. The main problem is that we cannot find a way of formulating the gist of Proposition 1, which looks to be material to a proof of the theorem, without referring to the spots s_1, \dots, s_{k+2} .

10 PGLD with Indirect Jumps

In this section, we introduce a variant of PGLD with indirect jump instructions. This variant is called PGLD_{ij}. However, preceding the introduction of PGLD_{ij}, we give a state-based description of the very simple family of services that constitute a register file of which the registers can contain natural numbers up to some bound. This register file will be used later on to describe the behaviour of PGLD_{ij} programs.

It is assumed that a fixed but arbitrary number I has been given, which is considered the number of registers available. It is also assumed that a fixed but arbitrary number N has been given, which is considered the greatest natural number that can be contained in a register.

The register file services accept the following methods:

- for each $i \in [0, I]$ and $n \in [0, N]$, a *register set method* $\text{set}:i:n$;
- for each $i \in [0, I]$ and $n \in [0, N]$, a *register test method* $\text{eq}:i:n$.

We write $\mathcal{M}_{\text{regs}}$ for the set $\{\text{set}:i:n, \text{eq}:i:n \mid i \in [0, I] \wedge n \in [0, N]\}$. It is assumed that $\mathcal{M}_{\text{regs}} \subseteq \mathcal{M}$.

The methods accepted by register file services can be explained as follows:

- $\text{set}:i:n$: the contents of register i becomes n and the reply is T;
- $\text{eq}:i:n$: if the contents of register i equals n , then nothing changes and the reply is T; otherwise nothing changes and the reply is F.

Let $s : [1, I] \rightarrow [0, N]$. Then we write Regs_s for the service with initial state s described by $S = ([1, I] \rightarrow [0, N]) \cup \{\uparrow\}$, where $\uparrow \notin [1, I] \rightarrow [0, N]$, and the functions eff , yld and act defined as follows ($i \in [0, I]$, $n \in [0, N]$, $\rho : [1, I] \rightarrow [0, N]$):

$$\begin{aligned}
\text{eff}(\text{set}:i:n, \rho) &= \rho \oplus [i \mapsto n] , \\
\text{eff}(\text{eq}:i:n, \rho) &= \rho , \\
\text{eff}(m, \rho) &= \uparrow && \text{if } m \notin \mathcal{M}_{\text{regs}} , \\
\text{eff}(m, \uparrow) &= \uparrow , \\
\text{yld}(\text{set}:i:n, \rho) &= \text{T} , \\
\text{yld}(\text{eq}:i:n, \rho) &= \text{T} && \text{if } \rho(i) = n , \\
\text{yld}(\text{eq}:i:n, \rho) &= \text{F} && \text{if } \rho(i) \neq n , \\
\text{yld}(m, \rho) &= \text{B} && \text{if } m \notin \mathcal{M}_{\text{regs}} , \\
\text{yld}(m, \uparrow) &= \text{B} , \\
\text{act}(m, \rho) &= \text{tau} , \\
\text{act}(m, \uparrow) &= \text{tau} .
\end{aligned}$$

We write $\text{Regs}_{\text{init}}$ for $\text{Regs}_{[1 \mapsto 0] \oplus \dots \oplus [I \mapsto 0]}$.

We pass on to PGLD with indirect jump instructions. In PGLD_{ij}, it is assumed that there is a fixed but arbitrary finite set of *foci* \mathcal{F} with $\text{regs} \in \mathcal{F}$ and a

fixed but arbitrary finite set of *methods* \mathcal{M} . Moreover, we adopt the assumptions made about register file services above. The set $\{f.m \mid f \in \mathcal{F} \setminus \{\text{regs}\}, m \in \mathcal{M}\}$ is taken as the set \mathfrak{A} of basic instructions.

PGLD_{ij} has the following primitive instructions:

- for each $a \in \mathfrak{A}$, a *plain basic instruction* a ;
- for each $a \in \mathfrak{A}$, a *positive test instruction* $+a$;
- for each $a \in \mathfrak{A}$, a *negative test instruction* $-a$;
- for each $l \in \mathbb{N}$, a *direct absolute jump instruction* $##l$;
- for each $i \in [1, I]$ and $n \in [1, N]$, a *register set instruction* $\text{set}:i:n$;
- for each $i \in [1, I]$, an *indirect absolute jump instruction* $i##i$.

PGLD_{ij} programs have the form $u_1 ; \dots ; u_k$, where u_1, \dots, u_k are primitive instructions of PGLD_{ij}.

The plain basic instructions, the positive test instructions, the negative test instructions, and the direct absolute jump instructions are as in PGLD. The effect of a register set instruction $\text{set}:i:n$ is that the content of register i becomes n . The effect of an indirect absolute jump instruction $i##i$ is that execution proceeds with the l -th instruction of the program concerned, where l is the content of register i . If $i##i$ is itself the l -th instruction, then deadlock occurs. If l equals 0 or l is greater than the length of the program, termination occurs. We stipulate that the content of each register is initially 0.

Like before, we define the meaning of PGLD_{ij} programs by means of a function pgldij2pgld from the set of all PGLD_{ij} programs to the set of all PGLD programs. This function is defined by

$$\begin{aligned} \text{pgldij2pgld}(u_1 ; \dots ; u_k) = & \\ & \psi(u_1) ; \dots ; \psi(u_k) ; ##0 ; ##0 ; \\ & +\text{regs.eq}:1:1 ; ##1 ; \dots ; +\text{regs.eq}:1:n ; ##n ; ##0 ; \\ & \vdots \\ & +\text{regs.eq}:I:1 ; ##1 ; \dots ; +\text{regs.eq}:I:n ; ##n ; ##0 , \end{aligned}$$

where $n = \min(k, N)$ and the auxiliary function ψ from the set of all primitive instructions of PGLD_{ij} to the set of all primitive instructions of PGLD is defined as follows:

$$\begin{aligned} \psi(a) &= a , \\ \psi(+a) &= +a , \\ \psi(-a) &= -a , \\ \psi(##l) &= ##l && \text{if } l \leq k , \\ \psi(##l) &= ##0 && \text{if } l > k , \\ \psi(\text{set}:i:n) &= \text{regs.set}:i:n , \\ \psi(i##i) &= ##l_i , \end{aligned}$$

and for each $i \in [1, I]$:

$$l_i = k + 3 + (2 \cdot \min(k, N) + 1) \cdot (i - 1) .$$

The idea is that each indirect absolute jump can be replaced by a direct absolute jump to the beginning of the instruction sequence

$+regs.eq:i:1 ; ##1 ; \dots ; +regs.eq:i:n ; ##n ; ##0 ,$

where i is the register concerned and $n = \min(k, N)$. The execution of this instruction sequence leads to the intended jump after the content of the register concerned has been found by a linear search. To enforce termination of the program after execution of its last instruction if the last instruction is a plain basic instruction, a positive test instruction or a negative test instruction, $##0 ; ##0$ is appended to $\psi(u_1) ; \dots ; \psi(u_k)$. Because the length of the translated program is greater than k , care is taken that there are no direct absolute jumps to instructions with a position greater than k . Obviously, the linear search for the content of a register can be replaced by a binary search.

Let P be a $PGLD_{ij}$ program. Then $pgldij2pgld(P)$ represents the meaning of P as a PGLD program. The intended behaviour of P is the behaviour of $pgldij2pgld(P)$ on interaction with a register file when abstracted from τ . That is, the *behaviour* of P , written $|P|_{PGLD_{ij}}$, is $\tau_{\tau}(|pgldij2pgld(P)|_{PGLD} /_{regs} Regs_{init})$.

A slightly different variant of PGLD with indirect jump instructions is introduced in [7] under the same name.

11 The Interpretation of $PGLD_{ij}$ Programs

In this section, we associate each $PGLD_{ij}$ program with a PGA program for constructing a representation of the $PGLD_{ij}$ program by a molecule and introduce a PGA program for interpreting $PGLD_{ij}$ programs represented by molecules. This amounts to enhancing the PGA programs given in Sections 8 and 9. $PGLD_{ij}$ programs without indirect absolute jump instructions are represented and interpreted exactly as before.

The idea of the enhancements is that:

- an atom is created for each of the registers;
- each atom that corresponds to a register is handled as if it concerns a direct jump instruction to the instruction with the content of the register as position;
- each atom that corresponds to an indirect jump instruction is handled as if it concerns a direct jump instruction to the direct jump instruction that takes the place of the register concerned;
- each register set instruction is handled as if it concerns an instruction for overwriting the direct jump instruction that takes the place of the register concerned.

We define the PGA programs for constructing representations of $PGLD_{ij}$ programs by means of a function $pgldij2md$ from the set of all $PGLD_{ij}$ programs to the set of all PGA programs. This function is defined by

$$\begin{aligned}
\text{pgldij2md}(u_1; \dots; u_k) = & \\
& f_1!; \dots; f_n!; m_1!; \dots; m_{n'}!; s_1!; \dots; s_{k+2}!; s'_1!; \dots; s'_I!; \\
& \rho_1(u_1); \dots; \rho_k(u_k); \\
& s_{k+1}/\text{stop}; s_{k+2}/\text{stop}; \\
& s'_1/\text{jmp}; s'_1.\text{jmp} = s_{k+2}; \dots; s'_I/\text{jmp}; s'_I.\text{jmp} = s_{k+2}; \\
& s = s_1; !,
\end{aligned}$$

where the auxiliary functions ρ_j from the set of all primitive instructions of PGLD_{ij} to the set of all PGA programs are defined as follows ($1 \leq j \leq k$):

$$\begin{aligned}
\rho_j(f.m) &= s_j/\text{focus}; s_j/\text{meth}; s_j/\text{pos}; s_j/\text{neg}; \\
&\quad s_j.\text{focus} = f; s_j.\text{meth} = m; s_j.\text{pos} = s_{j+1}; s_j.\text{neg} = s_{j+1}, \\
\rho_j(+f.m) &= s_j/\text{focus}; s_j/\text{meth}; s_j/\text{pos}; s_j/\text{neg}; \\
&\quad s_j.\text{focus} = f; s_j.\text{meth} = m; s_j.\text{pos} = s_{j+1}; s_j.\text{neg} = s_{j+2}, \\
\rho_j(-f.m) &= s_j/\text{focus}; s_j/\text{meth}; s_j/\text{pos}; s_j/\text{neg}; \\
&\quad s_j.\text{focus} = f; s_j.\text{meth} = m; s_j.\text{pos} = s_{j+2}; s_j.\text{neg} = s_{j+1}, \\
\rho_j(\#\#l) &= s_j/\text{jmp}; s_j.\text{jmp} = s_l && \text{if } 1 \leq l \leq k, \\
\rho_j(\#\#l) &= s_j/\text{stop} && \text{if } \neg(1 \leq l \leq k), \\
\rho_j(\text{set}:i:l) &= s_j/\text{reg}; s_j/\text{cont}; s_j/\text{nxt}; \\
&\quad s_j.\text{reg} = s'_i; s_j.\text{cont} = s_l; s_j.\text{nxt} = s_{j+1} && \text{if } 1 \leq l \leq k, \\
\rho_j(\text{set}:i:l) &= s_j/\text{reg}; s_j/\text{cont}; s_j/\text{nxt}; \\
&\quad s_j.\text{reg} = s'_i; s_j.\text{cont} = s_{k+2}; s_j.\text{nxt} = s_{j+1} && \text{if } \neg(1 \leq l \leq k), \\
\rho_j(i\#\#i) &= s_j/\text{jmp}; s_j.\text{jmp} = s'_i
\end{aligned}$$

and

- $f_1, \dots, f_n \in \mathcal{F}$ are the different foci that occur in $u_1; \dots; u_k$;
- $m_1, \dots, m_{n'} \in \mathcal{M}$ are the different methods that occur in $u_1; \dots; u_k$;
- $s, s_1, \dots, s_{k+2}, s'_1, \dots, s'_I \in \text{Spot} \setminus (\mathcal{F} \cup \mathcal{M})$.

The following is the PGA program for interpreting PGLD_{ij} programs represented by molecules:

$$\begin{aligned}
& (+s|\text{stop}; !; \\
& +s|\text{jmp}; \#16; \\
& +s|\text{reg}; \#9; \\
& u = s.\text{focus}; v = s.\text{meth}; +\%u.\%v; \#3; s = s.\text{neg}; \#9; s = s.\text{pos}; \#7; \\
& u = s.\text{reg}; v = s.\text{cont}; u.\text{jmp} = v; s = s.\text{nxt}; \#2; \\
& s = s.\text{jmp})^\omega,
\end{aligned}$$

where $u, v \in \text{Spot} \setminus (\mathcal{F} \cup \mathcal{M})$. Below, we write I' for this PGA program.

Theorem 2 below states rigorously that program I' interprets PGLD_{ij} programs correctly. In that theorem, other than in Theorem 1, we write $\mathcal{MD}_{\text{repr}(P)}$ for $|\text{pgldij2md}(P)| \bullet_{\text{md}} \mathcal{MD}_{\text{init}}$.

Theorem 2. *For all PGLD_{ij} programs P :*

$$|P|_{\text{PGLD}_{ij}} = \tau_{\text{tau}}(|I'| /_{\text{md}} \mathcal{MD}_{\text{repr}(P)}) .$$

Proof. The proof follows the same line as the proof of Theorem 1. This is possible because on interpretation any change of the state of the register file is reflected by a corresponding change of its molecular representation. \square

12 Conclusions

In this paper, we have considered the programming of an interpreter for a program notation that is close to existing assembly languages using PGA with the primitives of molecular dynamics as basic instructions. We have given PGA programs for constructing representations of the programs to be interpreted by molecules and a PGA program for interpreting those representations and we have shown that the latter PGA program does the interpretation correctly. We have experienced that, although primarily meant for explaining programming language features relating to the use of dynamic data structures, the collection of primitives of molecular dynamics in itself is suited to the programming wants concerned.

We observe that: (i) the program notation for which the presented interpreter has been designed belongs to the simplest program notations devised ever, (ii) it is hard to imagine that the programs to be interpreted can be represented by molecules in a way that is simpler than the way chosen for the presented interpreter, (iii) it is hard to conceive of an interpreter that is simpler than the presented interpreter. This means not at all that the design of the interpreter was simple. On the contrary, the design turned out to be disappointingly difficult. It happened that, owing to the quest for a simple interpreter, it was inescapable that the design was to a great extent a trial-and-error matter.

Dynamic data structures modelled using molecular dynamics can straightforwardly be implemented in programming languages ranging from PASCAL [15] to C# [11] through pointers or references, provided that fields are not added or removed dynamically. Using molecular dynamics, we need not be aware of the existence of the pointers used for linking data. The name molecular dynamics refers to the molecule metaphor used in the introduction. By that, there is no clue in the name itself to what it stands for. To remedy this defect, we suggest data linkage dynamics as an alternative name.

References

1. J. A. Bergstra and I. Bethke. Molecular dynamics. *Journal of Logic and Algebraic Programming*, 51:193–214, 2002.
2. J. A. Bergstra and I. Bethke. Polarized process algebra and program equivalence. In J. C. M. Baeten, J. K. Lenstra, J. Parrow, and G. J. Woeginger, editors, *Proceedings 30th ICALP*, volume 2719 of *Lecture Notes in Computer Science*, pages 1–21. Springer-Verlag, 2003.

3. J. A. Bergstra and M. E. Loots. Program algebra for sequential code. *Journal of Logic and Algebraic Programming*, 51(2):125–156, 2002.
4. J. A. Bergstra and C. A. Middelburg. Thread algebra with multi-level strategies. *Fundamenta Informaticae*, 71(2/3):153–182, 2006.
5. J. A. Bergstra and C. A. Middelburg. A thread calculus with molecular dynamics. Computer Science Report 06-24, Department of Mathematics and Computer Science, Eindhoven University of Technology, August 2006.
6. J. A. Bergstra and C. A. Middelburg. Instruction sequences with dynamically instantiated instructions. Electronic Report PRG0710, Programming Research Group, University of Amsterdam, November 2007.
7. J. A. Bergstra and C. A. Middelburg. Instruction sequences with indirect jumps. Electronic Report PRG0709, Programming Research Group, University of Amsterdam, November 2007.
8. J. A. Bergstra and C. A. Middelburg. Thread algebra for strategic interleaving. *Formal Aspects of Computing*, 19(4):445–474, 2007.
9. J. A. Bergstra and A. Ponse. Combining programs and state machines. *Journal of Logic and Algebraic Programming*, 51(2):175–192, 2002.
10. W. J. Fokkink. *Introduction to Process Algebra*. Texts in Theoretical Computer Science, An EATCS Series. Springer-Verlag, Berlin, 2000.
11. A. Hejlsberg, S. Wiltamuth, and P. Golde. *C# Language Specification*. Addison-Wesley, Reading, MA, 2003.
12. A. Ponse and M. B. van der Zwaag. An introduction to program and thread algebra. In A. Beckmann et al., editors, *CiE 2006*, volume 3988 of *Lecture Notes in Computer Science*, pages 445–458. Springer-Verlag, 2006.
13. D. Sannella and A. Tarlecki. Algebraic preliminaries. In E. Astesiano, H.-J. Krewski, and B. Krieg-Brückner, editors, *Algebraic Foundations of Systems Specification*, pages 13–30. Springer-Verlag, Berlin, 1999.
14. M. Wirsing. Algebraic specification. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 675–788. Elsevier, Amsterdam, 1990.
15. N. Wirth. The programming language PASCAL. *Acta Informatica*, 1(1):35–63, 1971.